ON THE FISHER-BEHRENS TEST - AN M.S. THESIS PROBLEM(S)

W. T. Federer

BU-627-M October 1977

Abstract

The Fisher-Behrens test has been discussed for the case of two unpaired means. The case involving a contrast of three or more means from an unpaired (completely randomized) design appears not to have been considered. Suggestions for possible extensions of the Fisher-Behrens test for k treatment means in a contrast are presented. Analogous suggestions are made for the studentized range statistic and the F-statistic as used in Scheffé's multiple comparisons procedure. It should be noted that this problem does not arise in paired (blocked) designs since the individual error variance for a contrast may be computed directly and the degrees of freedom for the error variance are known.

ON THE FISHER-BEHRENS TEST – AN M.S. THESIS PROBLEM(S)

W. T. Federer

BU-627-M

October 1977

1. Introduction

One of the most common techniques taught in Statistics methodology courses is Student's t-test for unpaired samples. Occasionally the problem of unequal variances in the two samples is considered. Procedures for using the statistic

$$t' = (\bar{y}_1-\bar{y}_2)/\sqrt{s_1^2/n_1 + s_2^2/n_2} = \bar{d}/s_{\bar{d}} \quad , \tag{1.1}$$

where $\bar{y}_i$ is the sample mean of n_i observations, s_i^2 is the sample variance from a normal population with variance σ_i^2 with $n_i - 1 = f_i$ degrees of freedom, have been studied by a number of authors under the topic denoted as the Fisher-Behrens test. The statistic in (1.1) does not follow Student's t-distribution except in special cases. Furthermore, nothing appears to have been said about this problem except for contrasts of two means. The problem of a linear contrast among k sample means from populations with different error variances σ_i^2 appears to have been entirely ignored.

Starting with an approximation to Student's t-distribution, statistics are proposed for

(i) an approximate t-statistic for a linear contrast of k means,

(ii) an approximate form for a studentized range statistic q used in connection with Tukey's experimentwise hsd multiple comparisons procedure, and

(iii) an approximate form for Scheffé's multiple comparisons procedure

when the k sample means in the contrast come from populations with differing error variances. Note that there are v treatments in the experiment and that

$2 \leq k \leq v$ means are involved in the contrast. The error rate in (i) is comparisonwise, while the error rates in (ii) and (iii) are experimentwise.

It is proposed that the 9 conjectures presented about (i), (ii), and (iii) above be investigated experimentally, using computer simulation methods. Some suggested sample sizes, treatment numbers, and error variances are presented for possible investigation. Since the total number of simulations could be large, the problem could be split up by number of treatments, by conjectures, or by statistic. It should be noted that one set of samples can be used to compute all statistics presented. Hence, it would be more efficient to compute all statistics on one set of samples. Likewise, several single degree of freedom contrast matrices are suggested. They could all be computed on one set of samples.

In the discussion a suggestion is made to extend these procedures to other multiple comparisons procedures. If these statistics are to be studied, it may be wise to include them in the computer program in order to utilize the same set of samples. In addition, the problem of degrees of freedom for a combined intrablock and interblock variance in lattice designs and of a combined form of error (a) and error (b) variances in a split plot design are discussed in relation to the conjectures.

2. A Proposed Extension of a Form of the Fisher-Behrens Test When There Are v Treatment Means

The following formula has appeared in textbook form for over thirty years:

$$t'_\alpha = \frac{w_1 t_{\alpha f_1} + w_2 t_{\alpha f_2}}{w_1 + w_2} \qquad (2.1)$$

where $w_i = s_i^2/n_i$, n_i = sample size for treatment i with mean $\bar{y}_i$, the sample is associated with variance s_i^2, f_i=degrees of freedom, i=1,2, and t_{α,f_i} is the tabulated value of Student's t for f_i degrees of freedom at the α-percentage point. Formula (2.1) appeared in Cochran and Cox's mimeographed notes in 1944 and may be found in Snedecor [1946, 4th edition] page 84, Cochran and Cox [1950, 1st edition] page 92, and in Federer [1955] page 94, for example.

Let us consider that we have a linear contrast of k of v sample means, say $\Sigma_{i=1}^{k} c_i \bar{y}_i$ where $\Sigma_{i=1}^{k} c_i = 0$ and where $|c_i| = 1$. Then, for the case where the error variances σ_i^2 in the v populations are unequal, we propose the following statistic as an approximation to the Student t distribution:

$$t'_\alpha = \sum_{i=1}^{k} w_i t_{\alpha f_i} \Big/ \sum_{i=1}^{k} w_i \tag{2.2}$$

where $t_{\alpha f_i}$ is the tabulated value of Student's t at the α-percentage point with f_i degrees of freedom and where $w_i = s_i^2/n_i$ = sample variance for treatment i divided by sample size.

<u>Conjecture 1</u>: t'_α from (2.2) is "approximately" Student's t with f degrees of freedom when f_i = f for all $i=1,2,\cdots,k \le v$ means in the contrast.

<u>Conjecture 2</u>: t'_α from (2.2) has "approximately" the correct percentage points for all k, n_i, f_i, σ_i^2, and α.

When the $|c_i| \neq 1$ in the linear contrast $\Sigma_{i=1}^{k} c_i \bar{y}_i$, we suggest the following:

$$t'_\alpha = \sum_{i=1}^{k} |c_i| w_i t_{\alpha f_i} \Big/ \sum_{i=1}^{k} |c_i| w_i \quad . \tag{2.3}$$

<u>Conjecture 3</u>: t'_α from (2.3) has "approximately" the correct percentage points for all c_i, f_i, k, n_i, σ_i^2, and α.

The orthogonal contrast matrices suggested are the orthogonal polynomial, the Helmert, and the Hadamard (when v = 4t). The treatment numbers suggested are 3, 4, 5, and 10 with perhaps other numbers included after the results for these numbers are obtained. Some suggested variance and sample sizes for v = 3 and 4 are given in Tables 2.1 and 2.2, respectively. It is suggested that 500-1000 samples be obtained for each set of v treatments in Tables 2.1 and 2.2. Perhaps 100 samples would suffice for each point, since there is a large number of points.

Sample Sizes	Error Variances								
	σ_1^2	σ_2^2	σ_3^2	σ_1^2	σ_2^2	σ_3^2	σ_1^2	σ_2^2	σ_3^2
	1	1.5	2	1	2	4	1	4	8
n_1, n_2, n_3	5	5	10	5	5	10	5	5	10
n_1, n_2, n_3	5	10	20	5	10	20	5	10	20
n_1, n_2, n_3	10	20	40	10	20	40	10	20	40
n_1, n_2, n_3	10	5	5	10	5	5	10	5	5
n_1, n_2, n_3	20	10	5	20	10	5	20	10	5
n_1, n_2, n_3	40	20	10	40	20	10	40	20	10
n_1, n_2, n_3	10	40	20	10	40	20	10	40	20
n_1, n_2, n_3	20	40	10	20	40	10	20	40	10
n_1, n_2, n_3	3	5	7	3	5	7	3	5	7
n_1, n_2, n_3	5	5	5	5	5	5	5	5	5
n_1, n_2, n_3	10	10	10	10	10	10	10	10	10
n_1, n_2, n_3	20	20	20	20	20	20	20	20	20
n_1, n_2, n_3	40	40	40	40	40	40	40	40	40

Table 2.1. Sample sizes and error variance sizes for v = 3 treatments.

Sample Sizes	Error Variance σ_1^2 1	σ_2^2 1.5	σ_3^2 2.0	σ_4^2 2.5	σ_1^2 1	σ_2^2 2	σ_3^2 3	σ_4^2 4	σ_1^2 1	σ_2^2 2	σ_3^2 4	σ_4^2 8
n_1, n_2, n_3, n_4	5	6	7	8	5	6	7	8	5	6	7	8
n_1, n_2, n_3, n_4	8	7	6	5	8	7	6	5	8	7	6	5
n_1, n_2, n_3, n_4	5	10	15	20	5	10	15	20	5	10	15	20
n_1, n_2, n_3, n_4	20	5	10	15	20	5	10	15	20	5	10	15
n_1, n_2, n_3, n_4	15	20	5	10	15	20	5	10	15	20	5	10
n_1, n_2, n_3, n_4	10	15	20	5	10	15	20	5	10	15	20	5
n_1, n_2, n_3, n_4	5	15	25	40	5	15	25	40	5	15	25	40
n_1, n_2, n_3, n_4	40	5	15	25	40	5	15	25	40	5	15	25
n_1, n_2, n_3, n_4	25	40	5	15	25	40	5	15	25	40	5	15
n_1, n_2, n_3, n_4	15	25	40	5	15	25	40	5	15	25	40	5
n_1, n_2, n_3, n_4	5	5	5	5	5	5	5	5	5	5	5	5
n_1, n_2, n_3, n_4	10	10	10	10	10	10	10	10	10	10	10	10
n_1, n_2, n_3, n_4	20	20	20	20	20	20	20	20	20	20	20	20
n_1, n_2, n_3, n_4	40	40	40	40	40	40	40	40	40	40	40	40

Table 2.2. Sample sizes and error variance sizes for $v = 4$ treatments.

Other treatment numbers to be investigated probably should be 5, 10, and 20, especially for equal sample sizes of say 3, 5, 10, 20, and 40. This study would be related to conjecture 1. One study could be related to equal sample sizes for the treatment sample means and variances for various treatment numbers and contrasts.

3. Studentized-Range Statistic With Unequal Variances

A straight forward extension of the results in the previous section produces formulas for the studentized range statistic as follows:

$$q'_{\alpha,v} = (w_1 q_{\alpha,v,f_1} + w_2 q_{\alpha,v,f_2})/(w_1+w_2) \tag{3.1}$$

$$q'_{\alpha,v} = \sum_{i=1}^{k} w_i q_{\alpha,v,f_i} \Big/ \sum_{i=1}^{k} w_i \tag{3.2}$$

$$q'_{\alpha,v} = \sum |c_i| w_i q_{\alpha,v,f_i} \Big/ \sum_{i=1}^{k} |c_i| w_i \tag{3.3}$$

where (3.1) is for a contrast of two means, (3.2) is a contrast of k means with the contrast coefficients being ± 1, and (3.3) is for a linear contrast among k of the v means when the c_i are not all equal to ± 1.

<u>Conjecture 4</u>: $q'_{\alpha,v}$ from (3.1) is a close approximation to $q_{\alpha,v,f}$ when all $f_i = f$.

<u>Conjecture 5</u>: $q'_{\alpha,v}$ from (3.2) is "approximately" $q_{\alpha,v,f}$ when all $f_i = f$.

<u>Conjecture 6</u>: $q'_{\alpha,v}$ from (3.2) and (3.3) has "approximately" the correct percentage points for all α, k, f_i, n_i, and σ_i^2 .

It is suggested that the same set of samples for v = 3, 4, 5, and perhaps other values be used to obtain experimental evidence for the "truth" of conjectures 4, 5, and 6.

4. <u>Scheffé's Statistic With Unequal Variances</u>

The proposed formulae for Scheffé's multiple comparisons procedure are given below. For a contrast of two means, we propose:

$$S'_{\alpha,v} = (w_1 S_{\alpha,v,f_1} + w_2 S_{\alpha,v,f_2})/(w_1+w_2) \tag{4.1}$$

where $S_{\alpha,v,f_i} = \sqrt{(v-1)F_\alpha(v-1,f_i)}$, $F_\alpha(v-1,f_i)$ is the tabulated value of

Snedecor's F at the α-percentage point with v - 1 degrees of freedom and f_i degrees of freedom for the sample error variance s_i^2, and $w_i = s_i^2/n_i$ as before. Consider a linear combination of k sample means, say $\Sigma_{i=1}^k c_i\bar{y}_i$ where $\Sigma_{i=1}^k c_i = 0$, where the samples are from populations having different variances. Let $S'_{\alpha,v}$ be computed as:

$$S'_{\alpha,v} = \sum_{i=1}^{k} |c_i| w_i S_{\alpha,v,f_i} \Big/ \sum_{i=1}^{k} |c_i| w_i \ . \tag{4.2}$$

Conjecture 7: $S'_{\alpha,v}$ from (4.1) is a close approximation to $S_{\alpha,v,f}$ when $f_1 = f_2$.

Conjecture 8: $S'_{\alpha,v}$ from (4.2) is a close approximation to $S_{\alpha,v,f}$ for all $f_i = f$ and all $2 \leq k \leq v$.

Conjecture 9: $S'_{\alpha,v}$ from (4.2) has approximately the correct percentage points for all k, α, f_i, n_i, and σ_i^2 .

5. Discussion

Two other statistics that come to mind are the esd and Dunnett's comparison with a control. The former has a per-experiment error rate of α, while the latter has an α-experimentwise error rate. These could be included with the lsd (Student's t), hsd, and Scheffé's procedures on the same set of samples. Duncan's multiple range test and the "short cut to allowances" or rsd procedures could also be investigated. However, the rsd procedure is known to be sensitive to differences in variances and hence probably should not be studied.

Another situation that could be investigated is the one in which $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_r^2$ but all other error variances are unequal, i.e., $\sigma_{v-r}^2 \neq \sigma_{v-r+1}^2 \neq \cdots \neq \sigma_v^2$. The statistic in formula (2.2), e.g., could be changed so that s_1^2 is the pooled variance for samples $1,2,\cdots,r$ with $f_1 = (n_1+n_2+\cdots+n_r-r)$ degrees of freedom and then proceed as suggested. Sometimes groups of treatments have the same variances within groups, but different variances from group to group. Again, the formulae are easily modified to accommodate this situation.

One further study with unequal variances pertains to combining intrablock and interblock information in certain nonorthogonal blocked designs. For example, let s_1^2 and s_2^2 be two mean squares with f_1 and f_2 degrees of freedom, respectively. A combined formula for the two variances has been given as (see Federer [1955], page 369):

$$E = s_1^2 \left\{ 1 + \frac{r}{(r-1)(k+1)} - \frac{rs_1^2}{(r-1)(k+1)s_2^2} \right\} \tag{5.1}$$

where each of the v treatments is repeated r times and $k < v$ is the block size for n-ple lattice designs; $f_1 = (r-1)(v-1) - f_2$ and $f_2 = r(k-1)$. It is contended by Paul Meier (Ph.D. Thesis, Johns Hopkins under J. W. Tukey) that E is approximately distributed as chi-square with $f_1 + f_2$ degrees of freedom. He reached this conclusion because the first four moments of E compared closely to the first four moments of a chi-square with $f_1 + f_2$ degrees of freedom. The author has contended that the degrees of freedom should be between f_1 and f_2 and the fact that the first four moments of E compare with the first four moments of a chi-squares is irrelevant, since E is not a chi-square but a chi-square times a related F. Could the above simulations throw light on this problem?

Another method for using combined variances is given by formulae (X-6) and (X-19) in Federer [1955]. The degrees of freedom for these combined mean squares is unknown (at least to the author). Here again one should study the article by J. Taylor in Biometrika 37:443-444, 1950, the other references listed in Federer [1955], page 279, and perhaps consider simulation.

D. S. Robson has performed a number of simulations for the F-statistic in Spring 1976. Anyone contemplating a simulation is encouraged to discuss their proposed problem with him first. Other statisticians should also be consulted.

One additional item to consider is how the degrees of freedom x in $t'_{\alpha x}$ varies with respect to α. To illustrate, let

$$t'_{05} = \frac{2.145\left\{\frac{959.4}{15}\right\} + 2.201\left\{\frac{270.1}{12}\right\}}{\frac{959.4}{15} + \frac{270.1}{12}} = 2.16$$

and

$$t'_{01} = \frac{2.977\left\{\frac{959.4}{15}\right\} + 3.106\left\{\frac{270.1}{12}\right\}}{\frac{959.4}{15} + \frac{270.1}{12}} = 3.01 \quad .$$

$t'_{05} = 2.16$ corresponds to $t_{05,13}$ and $t'_{01} = 3.01$ corresponds to $t_{05,11.5}$. Now does t'_{α} approach the $t_{\alpha f_i}$ for f_i a minimum as α approaches zero? Does it approach the t_{α,f_i} for maximum f_i as α approaches 50%? A study of this relationship may prove profitable; at least, one could establish bounds for various cases.

From the simulation one can prepare tables for approximate percentage points for t'_{α}, $q'_{\alpha,v}$, and $S'_{\alpha,v}$. A high degree polynomial can be used to fit a curve through all points and thus obtain more accurate percentage points than otherwise obtainable.

6. Literature Cited

1. Cochran, W. G. and Cox, G. M. [1950]. Experimental Designs, 1st edition (2nd edition 1957), Wiley and Sons, Inc., New York.

2. Federer, W. T. [1955]. Experimental Design., Macmillan, New York.

3. Snedecor, G. W. [1946]. Statistical Methods, 4th edition, Iowa State College Press, Ames, Iowa.

APPENDIX TO BU-627-M

An example illustrating the statistical computations for the simulation study proposed is presented below. The first 15 normal deviates from the 100,000 random normal deviates published by The Rand Corporation were used for the example below. v = 3 treatments with r = 5 replicates each were used. The data and computations are:

15 normal deviates	add 5 to keep nos. positive	
-1.276	3.724	$\bar{y}_{1.} = 23.227/5 = 4.6454 = 4.65$
-0.318	4.682	$s_1^2 = 2.43$
-1.377	3.623	$n_1 = 5$
2.334	7.334	$f_1 = 1,1,1$
-1.136	3.864	
0.414	5.414	$\bar{y}_{2.} = 26.592/5 = 5.39$
-0.494	4.506	$s_2^2 = 0.32$
1.048	6.048	$n_2 = 5$
0.347	5.347	$f_2 = 1.5,2,4$
0.637	5.637	
2.172	7.172	$\bar{y}_{3.} = 30.816/5 = 6.16$
-1.185	3.815	$s_3^2 = 2.19$
0.972	5.972	$n_3 = 5$
1.210	6.210	$f_3 = 2,4,8$
2.647	7.647	

$f_1 = 1,\ f_2 = 1.5,\ f_3 = 2$

variance of a difference is $\sqrt{\dfrac{s_i^2 f_i}{n_i} + \dfrac{s_j^2 f_j}{n_j}} = \sqrt{\frac{1}{5}(s_i^2 f_i + s_j^2 f_j)}$

$$t = \frac{\bar{y}_i - \bar{y}_j}{\sqrt{\frac{1}{5}(f_i s_i^2 + f_j s_j^2)}} \quad \text{for all pairs } ij$$

$$t_{12} = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{1}{5}\left(2.43 + 1.5(0.32)\right)}} = \frac{-0.74}{0.763} = -0.970$$

$$t_{13} = \frac{\bar{y}_1 - \bar{y}_3}{\sqrt{\frac{1}{5}\left(2.43 + 2(2.19)\right)}} = \frac{-1.51}{1.167} = -1.294$$

$$t_{23} = \frac{\bar{y}_2 - \bar{y}_3}{\sqrt{\frac{1}{5}\left(1.5(0.32) + 2(2.19)\right)}} = \frac{-0.77}{0.986} = -0.781$$

$f_1 = 1,\ f_2 = 2,\ f_3 = 4$

$$t_{12} = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{1}{5}\left(2.43 + 2(0.32)\right)}} = \frac{-0.74}{0.784} = -0.944$$

$$t_{13} = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{1}{5}\left(2.43 + 4(2.19)\right)}} = \frac{-1.51}{1.496} = -1.009$$

$$t_{23} = \frac{\bar{y}_2 - \bar{y}_3}{\sqrt{\frac{1}{5}\left(2(0.32) + 4(2.19)\right)}} = \frac{-.77}{1.371} = -0.562$$

$f_1 = 1,\ f_2 = 4,\ f_3 = 8$

$$t_{12} = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{1}{5}(s_1^2 + 4s_2^2)}} = \frac{-0.74}{0.861} = -0.856$$

$$t_{13} = \frac{\bar{y}_1 - \bar{y}_3}{\sqrt{\frac{1}{5}(s_1^2 + 8s_3^2)}} = \frac{-1.51}{1.997} = -0.756$$

$$t_{23} = \frac{\bar{y}_2 - \bar{y}_3}{\sqrt{\frac{1}{5}(4s_2^2 + 8s_3^2)}} = \frac{-0.77}{1.939} = -0.397$$

Now consider the differences utilizing all $v = 3$ means as follows:

$$-2\bar{y}_1 + \bar{y}_2 + \bar{y}_3 = +2.25$$
$$\bar{y}_1 - 2\bar{y}_2 + \bar{y}_3 = 0.03$$
$$\bar{y}_1 + \bar{y}_2 - 2\bar{y}_3 = -2.28$$

$f_1 = 1,\ f_2 = 1.5,\ f_3 = 2$

$$t_{2+3 \text{ vs. } 1} = \frac{2.25}{\sqrt{\frac{1}{5}(4s_1^2 + 1.5s_2^2 + 2s_3^2)}} = \frac{2.25}{1.708}$$

$$t_{1+3 \text{ vs. } 2} = \frac{0.03}{\sqrt{\frac{1}{5}(s_1^2 + 4(1.5)s_2^2 + 2s_3^2)}} = \frac{0.03}{1.321}$$

$$t_{1+2 \text{ vs. } 3} = \frac{-2.28}{\sqrt{\frac{1}{5}(s_1^2 + 1.5s_2^2 + 4(2)s_3^2)}} = \frac{-2.28}{2.021}$$

$f_1 = 1,\ f_2 = 2,\ f_3 = 4$

$$t_{2+3 \text{ vs. } 1} = \frac{2.25}{\sqrt{\frac{1}{5}(4s_1^2 + 2s_2^2 + 4s_3^2)}} =$$

$$t_{1+3 \text{ vs. } 2} = \frac{0.03}{\sqrt{\frac{1}{5}(s_1^2 + 2(4)s_2^2 + 4s_3^2)}} =$$

$$t_{1+2 \text{ vs. } 3} = \frac{-2.28}{\sqrt{\frac{1}{5}(s_1^2 + 2s_2^2 + 4(4)s_3^2)}} =$$

$f_1 = 1,\ f_2 = 4,\ f_3 = 8$

$$t_{2+3 \text{ vs. } 1} = \frac{2.25}{\sqrt{\frac{1}{5}(4s_1^2 + 4s_2^2 + 8s_3^2)}} =$$

$$t_{1+3 \text{ vs. } 2} = \frac{0.03}{\sqrt{\frac{1}{5}(s_1^2 + 4(4)s_2^2 + 8s_3^2)}} =$$

$$t_{1+2 \text{ vs. } 3} = \frac{-2.28}{\sqrt{\frac{1}{5}(s_1^2 + 4s_2^2 + 4(8)s_3^2)}} =$$

Obtain 1000 (2000?) samples for each of the above 12 t-statistics. If the conjectures are true, these statistics should follow a t-statistic with 5 - 1 = 4 degrees of freedom. If not, try to determine if they have a t-distribution with x degrees of freedom. Since we are interested in the tails of the distribution, 2000 samples may not be sufficient.

On the same sample one can compute studentized-range statistics. The range statistic q for v treatments is:

$$q = \frac{\bar{y}_{\text{maximum}} - \bar{y}_{\text{minimum}}}{\sqrt{\text{Variance of sample mean}}} .$$

Since the means have different variances, we could consider the standard error of a sample mean to be the square root of an arithmetic average of the two variances, which is:

$$\sqrt{[V(\bar{y}_{max.}) + V(\bar{y}_{min.})]/2}$$

or to be the geometric mean of the two variances, which is:

$$\left(\sqrt{V(\bar{y}_{max.})V(\bar{y}_{min.})}\right)^{\frac{1}{2}} .$$

Perhaps both forms of the variance should be used to compute the q statistic. In addition it is suggested that q statistics be computed for largest mean minus smallest mean, as follows:

$$q = \frac{\bar{y}_{max.} - \bar{y}_{min.}}{\sqrt{\frac{1}{2}[V(\bar{y}_{max.}) + V(\bar{y}_{min.})]}} \quad \text{and} \quad q = \frac{\bar{y}_{max.} - \bar{y}_{min.}}{\left(\sqrt{V(\bar{y}_{max.})V(\bar{y}_{min.})}\right)^{\frac{1}{2}}}$$

where $\bar{y}_{max.}$ is the largest sample mean in the set of v means for treatment i with variance $f_i s_i^2 = V(\bar{y}_i)$ and $\bar{y}_{min.}$ is the smallest mean in the set for treatment j and has variance $f_j s_j^2$. For the example

$$q = \frac{\bar{y}_{3.} - \bar{y}_{1.}}{\sqrt{(2s_3^2 + s_1^2)/2}} = \frac{1.51}{\sqrt{\frac{1}{2}\left(2(2.19) + 2.43\right)}} \quad \text{and} \quad q = \frac{\bar{y}_{3.} - \bar{y}_{1.}}{\left(\sqrt{2s_3^2 s_1^2}\right)^{\frac{1}{2}}} = \frac{1.51}{\left(\sqrt{2(2.19)(2.43)}\right)^{\frac{1}{2}}} .$$

One should also compute all other q statistics for other pairs of means. For the above example, the second largest difference of means is between 2 and 3:

$$q = \frac{6.16 - 5.39}{\left(\sqrt{(1.5)(0.32)(2)(2.19)}\right)^{\frac{1}{2}}} \quad \text{or} \quad q = \frac{6.16 - 5.39}{\sqrt{\frac{1}{2}[1.5(0.32) + 2(2.19)]}} .$$

$$q = \frac{5.39 - 4.65}{\left(\sqrt{2(0.32)(2.43)}\right)^{\frac{1}{2}}} \quad \text{or} \quad q = \frac{5.39 - 4.65}{\sqrt{\frac{1}{2}[2(0.32) + 2.43]}} .$$

Using all three sets of variances, we would obtain 18 such q-statistics; 9 using an arithmetic average of the variances and 9 using a geometric average of the variances. The distribution of each of the 18 should be compared with the distribution of a q-statistic with 4 degrees of freedom.

Scheffé's multiple comparisons procedure can be put in the form of an F-statistic as follows:

$$F = \frac{(\bar{y}_i - \bar{y}_j)^2}{\frac{1}{v-1}\left(\frac{s_i^2 f_i}{n_i} + \frac{s_j^2 f_j}{n_j}\right)} = \frac{(\bar{y}_i - \bar{y}_j)^2}{\frac{1}{3-1}\left(\frac{1}{5}\right)\left(f_i s_i^2 + f_j s_j^2\right)} \quad \text{for our case}$$

and

$$F = \frac{(\bar{y}_i + \bar{y}_j - 2\bar{y}_k)^2}{\frac{1}{(v-1)}\left(\frac{s_i^2 f_i}{n_i} + \frac{s_j^2 f_j}{n_j} + \frac{4 s_k^2 f_k}{n_k}\right)}$$

$$= \frac{(\bar{y}_i + \bar{y}_j - 2\bar{y}_k)^2}{\frac{1}{3-1}\left(\frac{1}{5}\right)\left(f_i s_i^2 + f_j s_j^2 + 4 f_k s_k^2\right)} \quad \text{for our case.}$$

One can determine maximum F and compare with F(v-1, n-1) distribution. Here again one should keep track of largest F, second largest, etc. to the smallest, as well as maximum difference of means either in pairs or triplets, the second largest, etc. to the smallest. The frequency distribution for each can be obtained.